

## BIROn - Birkbeck Institutional Research Online

Brown, C.R.H. and Berggren, Nick and Forster, S. (2020) Testing a goal-driven account of involuntary attentional capture by threat. *Emotion* 20 (4), pp. 572-589. ISSN 1528-3542.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/28023/>

*Usage Guidelines:*

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html>  
contact [lib-eprints@bbk.ac.uk](mailto:lib-eprints@bbk.ac.uk).

or alternatively

RUNNING HEAD: TESTING GOAL-DRIVEN CAPTURE BY THREAT

Testing a goal-driven account of involuntary attentional capture by threat

Chris R.H. Brown<sup>+</sup>\*, Nick Berggren<sup>#</sup>, & Sophie Forster<sup>+</sup>\*

+ School of Psychology, University of Sussex, UK

# Birkbeck, University of London, UK

Total word count (excluding abstract): 11464

\*Corresponding authors

School of Psychology, University of Sussex, Falmer, BN1 9QH, United Kingdom

Email: [chris.rh.brown13@gmail.com](mailto:chris.rh.brown13@gmail.com) or [s.forster@sussex.ac.uk](mailto:s.forster@sussex.ac.uk)

**Authors' note:**

All tasks, unlicensed images sourced or created for this investigation, as well as data and analysis scripts are available via the Open Science Framework: [osf.io/mr5yk](https://osf.io/mr5yk).

**Conflict of interests**

The authors declare that there were no conflicts of interest to report.

**Role of funding**

The authors declare that there was no external funding to report.

### **Abstract**

Attention has long been characterised within prominent models as reflecting a competition between goal-driven and stimulus-driven processes. It remains unclear, however, how involuntary attentional capture by affective stimuli, such as threat-laden content, fits into such models. While such effects were traditionally held to reflect stimulus-driven processes, recent research has increasingly implicated a critical role of goal-driven processes. Here we test an alternative goal-driven account of involuntary attentional capture by threat, using an experimental manipulation of goal-driven attention. To this end we combined the classic ‘contingent capture’ and ‘emotion-induced blink’ (EIB) paradigms in an RSVP task with both positive or threatening target search goals. Across six experiments, positive and threat distractors were presented in peripheral, parafoveal, and central locations. Across all distractor locations, we found that involuntary attentional capture by irrelevant threatening distractors could be induced via the adoption of a search goal for a threatening category; adopting a goal for a positive category conversely led to capture only by positive stimuli. Our findings provide direct experimental evidence for a causal role of voluntary goals in involuntary capture by irrelevant threat stimuli, and hence demonstrate the plausibility of a top-down account of this phenomenon. We discuss the implications of these findings in relation to current cognitive models of attention and clinical disorders.

Abstract word count: 211

Key words: threat; attentional bias; selective attention; goal-driven attention; top-down attention

In daily life, selective attention allows us to make sense of an otherwise overwhelming volume of perceptual input, prioritising the processing of stimuli that are in some way flagged as important (e.g. the words on a computer screen, or a voice over a phone), over stimuli that may have less importance (e.g. email pop-ups, a colleague passing by, or the tactile sensation of sitting in a chair). Some stimuli are selected intentionally, in line with our current goals, while others are selected in an involuntary manner such as stimuli with high perceptual salience. Prominent models of attention account well for the above examples within frameworks involving two key drivers of attention: a goal-driven ‘endogenous’ mechanism which directs attention in a strategic top-down manner, and a stimulus-driven ‘exogenous’ mechanism which directs attention in an involuntary manner to perceptually salient stimuli (Desimone & Duncan, 1995; Theeuwes, 1994; 2010; Itti & Koch, 2001; Buschman & Miller, 2007; Parkhurst, Law & Niebur, 2002; Turatto & Galfano, 2000). However, in daily life, stimuli may also catch our attention involuntarily, not due to low-level perceptual salience, but rather due to the affective content; for example, being associated with a potential threat (e.g. a spider on the office wall). It is not readily apparent how this form of attentional capture can be accommodated within the goal-driven and stimulus-driven dichotomy which is included in these models of attention; indeed, this problem has led to calls for theoretical revisions involving a third driver of attention (cf. Awh, Belopolsky & Theeuwes, 2012).

The omission of affective stimuli from these models of selective attention may have arisen, in part, because the experimental paradigms which shaped these theories involved simple, affectively neutral stimuli, such as basic geometric shapes or letters (e.g. Theeuwes, 1992; Yantis, 1993; Folk, Remington & Johnson, 1992). Such situations limit the likely influences on attention to two factors: the task instructions influencing goal-driven attention, and the perceptual stimulus-driven salience of the components of the stimulus display. On the

other hand, over the last several decades a rich literature has amassed regarding the study of attentional capture by affective stimuli such as threat. This research was initially conducted largely with relation to anxiety-related attentional biases. However, recent work has begun to investigate threat relevant stimuli more broadly within the framework of prominent selective attention models (e.g. Schmidt, Belopolsky & Theeuwes, 2015; Notebaert, Crombez, Van Damme, De Houwer & Theeuwes, 2011). Based on this, there is considerable empirical evidence to suggest that affective stimuli, including threat, can capture attention in a seemingly involuntary manner (see Carretié, 2014; Vuilleumier, 2005; 2015 for reviews).

Within the framework of the goal-driven/stimulus-driven dichotomy, involuntary selection of affective stimuli, such as threat, might initially be presumed to be stimulus-driven due to the apparent ability of stimulus-content to override task goals. Indeed, the stimulus-driven view of attentional capture has traditionally been the prevalent interpretation of such effects (Le Doux, 1995; 1998; Öhman, 1992; 2005; Öhman & Mineka, 2001). Influential theories of attention in anxiety, for example, have suggested that the attentional bias to threat stimuli is due to their learnt salience increasing bottom-up perceptual input; with highly anxious individuals being more sensitive to certain categories of stimulus-specific salience (Bishop, 2007; 2009; Cisler & Koster, 2010; Mogg & Bradley, 2016).

Conversely, the stimulus-driven account has recently been challenged by compelling evidence that attentional capture by threat may not be unconditional, and may only occur when the threatening stimulus is task-relevant (Hahn & Gronlund, 2007; Everaert, Spruyt & De Houwer, 2013; Everaert, Spruyt, Rossi, Pourtois & De Houwer, 2014; Vromen, Lipp & Remington, 2016; Stein, Zwickel, Ritter, Kitzmantel & Schneider, 2009; Vogt, De Houwer, Crombez, & Van Damme, 2013; Lichtenstein-Vidne, Henik & Safadi, 2012; Van Dillen, Lakens & Van Den Bos, 2011; for reviews of top-down factors in attention and perception of threat see Mohanty & Sussman, 2013; Sussman, Jin & Mohanty, 2016). For example, Stein et

al. (2009) examined the degree to which the emotional expressions of face search targets would heighten the attentional blink (AB) effect (i.e. impeding detection of a second target). Fearful versus neutral targets only produced heightened AB when the target response involved classifying the emotion of the face – when the response was non-emotional (male versus female) the same stimuli showed no difference in the AB effect for fearful versus neutral targets. Such findings appear more consistent with a goal-driven, rather than stimulus-driven mechanism, which would prioritise emotional stimuli with greater relevance to the current task goals.

It is intuitively plausible that individuals might commonly adopt top-down goals which prioritise threat detection. Such goals would be adaptive, allowing individuals to avoid potentially harmful outcomes by enhancing the ability to detect potential threats in the environment when experiencing fear or apprehension. The possibility that threat detection goals influence attention indeed accords well with the observation that biases in attention towards threatening stimuli are a hallmark symptom of anxiety (Cisler & Koster, 2010; Yiend, 2010). However, how could it be that voluntary goals for threat cause seemingly involuntary attentional capture?

One possibility is that in some experimental demonstrations of attentional capture by threat, the effect might in fact not reflect involuntary capture at all, but rather voluntary allocation of attention to the supposed ‘distractors’. The paradigms most commonly used to demonstrate attentional capture by threat (e.g. dot-probe, visual search, or emotional Stroop) present the threatening stimulus in a location that must be attended in order to successfully perform the task, often with no obvious performance cost to attending the images (although see Grimshaw, Kranz, Carmel, Moody & Devue, 2018, for a recent exception). It might, therefore, be argued that in such cases participants are simply failing to follow the instruction

to ignore the threat stimuli because it is not possible to both attend to a location and suppress the visual information presented there.

An alternative possibility, however, is that attentional capture by threat may be truly involuntary but nevertheless driven by top-down goals. While this may appear paradoxical, there is considerable evidence that involuntary attentional capture can in fact occur as an unintended consequence of voluntary goal-driven attention (Folk et al., 1992; Folk, Leber & Egeth, 2002; 2008; LeBlanc, Prime & Jolicoer, 2008). Compelling evidence from the ‘contingent capture’ literature suggests that when goal-driven attention is directed to a particular type of stimulus (e.g. a particular colour, shape, or even semantic category), any stimulus which matches the features which are currently being searched for may capture attention, even if they are in some way known to be irrelevant to the task (e.g. being presented in a task-irrelevant location). For example, when instructed to search for a letter of a particular colour in an RSVP stream of other coloured letters, an irrelevant peripherally presented distractor, which shares the target colour, captures attention and results in participants being unable to identify a subsequent coloured target (i.e. producing an attentional blink; Raymond, Shapiro & Arnell, 1992). Importantly, equally salient coloured distractors which do not share the specified target colour do not capture attention. In other words, participants searching for a green letter are typically distracted by peripheral green distractors but not red distractors, while participants searching for a red letter are distracted by red but not green distractors (Folk et al., 2002; 2008; Leblanc, Prime & Jolicoer, 2008).

The phenomenon of ‘contingent capture’ has recently been found to extend beyond low-level visual features to broadly defined goals, such as a conceptual category (e.g. office supplies or cars, Wyble, Folk & Potter, 2013; Reeder, van Zoest & Peelen, 2015). Critically, contingent capture occurs involuntarily, even though participants know that the stimulus is irrelevant to their current task goals (e.g. because it is presented in an irrelevant location in

which the task-relevant stimuli never appear). Could it be, then, that attentional capture by threat could occur as a form of affective contingent capture?

Here we provide the first direct test for a causal role of top-down goals in involuntary attentional capture by affective stimuli. Note that previous manipulations of relevance have only indirectly manipulated goal-driven attention. For example, Stein and colleagues' manipulation of response settings only indirectly manipulates goal-driven selection, as participants could conceivably have adopted a goal to select faces across all conditions and activated the response set post-selection. Other studies have manipulated affective distractor relevance by varying the affective content of search targets, but in contexts in which the instructed search goals are identical across conditions. For example, Lichtenstein-Vidne et al. (2012) found that during a task involving judging the location of centrally presented images, peripheral emotional images slowed responses only when the target images were also emotional, and not when the targets were neutral. In this paradigm, however, the emotional content was irrelevant to the instructed target selection criteria (an image presented above or below fixation), which in both conditions were based solely on location. While it appears plausible that participants may have elected to adopt a narrower goal than was necessary to perform the task, as goals were neither directly manipulated nor measured it remains unclear whether such relevance effects are goal-driven. Indeed, certain manipulations of relevance have been argued to affect stimulus-driven mechanisms (cf. Sui & Humphreys, 2015).

We posit that a direct test of whether goal-driven mechanisms can play a causal role in involuntary attentional capture by affective stimuli must meet the following conditions. First, the overlap between the capturing stimulus and goal-driven attentional settings must be manipulated directly by changing the criteria for selection (e.g. the search goal), such that a task can only be completed by adopting this goal. Second, attentional capture can only be assumed to be entirely involuntary if the task does not require any voluntary allocation of



attention to the capturing stimulus. Any demonstrations of attentional bias to a stimulus which participants are asked to search for (i.e. a search target), or to a stimulus presented in a potential target location (which necessarily requires some allocation of attention, cf. Forster, 2013), could potentially reflect the affective enhancement of voluntary attention rather than a truly involuntary attentional process; the affective content amplifying attention after it has been directed towards the stimulus, rather than causing the attentional capture itself. To our knowledge no prior study meets both of these criteria.

In order to directly test the goal-driven hypothesis, we therefore designed a task that would allow us to experimentally manipulate the participants' goals, and measure the effect of this manipulation on involuntary attentional capture by threat associated stimuli which appeared in task-irrelevant locations. To this end we fused the contingent capture paradigm (e.g. Folk et al., 2002; 2008; Wyble et al., 2013) with a well-established measure of attentional capture by threat: the emotion-induced blindness (EIB) paradigm, in which participants must respond to a target presented in an RSVP stream whilst ignoring an affective distractor presented beforehand in the same stream (e.g. Most, Chun, Widders & Zald, 2005; Smith, Most, Newsome & Zald, 2006; Zheng, Wang & Luo, 2015; Kennedy, Rawding, Most & Hoffman, 2014; Kennedy, Pearson, Sutton, Beesely & Most, 2018; Singh & Sunny, 2017; see McHugo, Olatunji & Zald, 2013 for review). Specifically, in our task, participants were instructed to search for a target stimulus defined by its affective category (e.g. positive or threatening), presented in a central RSVP stream while ignoring peripherally presented distractor images which were either positive, threatening, or neutral. Importantly, the target never appeared in the peripheral distractor locations meaning that it was never necessary to allocate voluntary attention to these locations, hence the distractor locations can be considered task-irrelevant. Our emotionally laden targets and distractors consisted of stimuli that have been widely used in the affective attentional bias literature: animals and

emotional faces (e.g. Lipp & Derakhshan, 2005; Fox, Russo & Dutton, 2002; Öhman, Flykt & Esteves, 2001; LoBue & Rakison, 2013). Goal-driven attentional capture would be reflected by greater interference from the threat distractors when these are congruent, versus incongruent, with the current search goal. Such a finding would demonstrate that involuntary attentional capture by threat could be plausibly accommodated within theories of attentional capture under the umbrella of the goal-driven mechanism.

## Experiment 1

### Methods

**Participants.** Twenty participants were initially recruited, though one participant was identified as an outlier and excluded due to their accuracy being 3 SDs below the group mean (16 females, 3 males; Age:  $M = 22.37$ ,  $SD = 3$ ). Our sample size of  $n = 19$  was selected on the basis of matching that used in the most similar prior demonstration of goal-driven capture (Wyble et al., 2013, Experiment 2). A power analysis using G\*power software indicated that this sample size would afford sensitivity to detect effect sizes above  $dz = .69$  with power of  $\beta = .80$  and an  $\alpha$  of .05 (Faul, Erdfelder, Lang, & Buchner, 2007). Hence, we were well powered to detect effects half the size of those observed by published demonstrations of goal-driven capture with non-affective stimuli; specifically, those comparing task-irrelevant goal-congruent and goal-incongruent distractors within an RSVP task with similar timings and settings to our design (Wyble et al., 2013: mean  $dz = 1.65$ ,  $SD = .38$ ; Folk et al., 2002: mean  $dz = 1.33$ ,  $SD = .25$ ).

Participants were recruited through the University of Sussex subject pool via an online advert. They were remunerated with course credits or a small cash payment. Ethical approval for this experiment, and all subsequent experiments, were granted by the University of Sussex Sciences and Technology Cross-Schools Research Ethics Committee.

**Stimuli.** The neutral animal stimuli were a range of animal images sourced from Google images. In total, 391 images of individual animals, without any other salient objects in the scene, were initially selected. The images were all resized to 300×200 pixels and all writing was removed. These images were rated in a pilot study by 36 participants using a ten-point Likert scale measuring how threatening, cute, positive and negative they were. To select the most neutral animals, a composite affect score was created by averaging these four scales together. The 280 images which were rated lowest on this measure were selected for neutral stimuli (Affect score:  $M = 3.23$ ,  $SD = .46$ , highest score = 4.01; highest positive score = 6.31; highest cute score = 6.83; highest negative score = 4.08; highest threat score = 3.58). From these 280 images we removed images which contained features which could be mistaken for part of the target set. For instance, many images of elephants, walruses, and water buffalo were removed because their horns and tusks could be mistaken for bared teeth. These images were replaced by 35 images of animals which were similar to those ranked in the lowest 150 images on overall affect (e.g. fish, birds, farm animals).

The threatening and cute animal images were partly selected from the International Affective Picture System (IAPS; Lang, Bradley & Cuthbert, 1997), but in order to provide a greater number of distinct threatening and cute animal images (reducing potential habituation effects) the IAPs images were supplemented with images from Google images. These latter images were selected based on their similarity to cute and threatening animals in the IAPS database; cute animals were usually pets or infant animals, whilst the threatening animals were either predators in attack positions or snakes and spiders. Based on these criteria we collected twelve target images and twelve different distractor images for the cute and threatening animal categories. The 24 images used in the threatening animal category (12 targets and 12 distractors) consisted of six different animals: spiders, lions, tigers, snakes, sharks, and crocodiles. For the cute category targets and distractors were comprised of six

different cute animals: kittens, puppies, pandas, red pandas, ducklings, and rabbits. Again, twelve images appeared as targets and twelve different images as distractors. For both cute and threatening categories, all six types of animals appeared as both targets and distractors, but not the individual images. To validate the images, arousal and valence ratings were collected again from participants in Experiments 3a, 3b, 4, and 5 (see Table 2) which confirmed that threat images were considered to have negative valence and be highly arousing. All unlicensed images and their ratings are available online via the Open Science Framework (link: [osf.io/mr5yk](https://osf.io/mr5yk)).

The images were presented using E-prime 2.0 on a 16inch Dell monitor with a screen resolution of 800×600 and refresh rate of 60Hz (Psychology Software Tools, Inc., 2012). The experiment was conducted in a dimly lit room. Participants viewed the screen from 59cm away, and this distance was kept constant by using a chin rest. All images in the central RSVP stream measured  $6^{\circ} \times 4.02^{\circ}$ . The distractors measured,  $8.09^{\circ} \times 5.35^{\circ}$ , these were larger relative to the central target due to visual acuity being poorer at peripheral locations. On every trial, the distractors were presented above and below the central RSVP stream with a gap of  $.5^{\circ}$  separation from the target. Trials were controlled so the specific animal presented as a distractor was never the same as the target animal.

**Procedure.** Figure 1 presents an example trial sequence in the experimental task. Participants were given the following instructions at the start of the task: “You will be shown several images of animals in quick succession. You must look out for either a 'cute' (e.g. baby or pet) or 'threatening' animal (e.g. predator or poisonous). You will be instructed which type of animal you are looking for before each trial. At the end of each trial you must write out the name of the cute/threatening animal using the keyboard. The target image will always appear in the centre of the screen. Occasionally two other images will appear at the top and bottom of the screen, you must ignore these images.”. Search goal reminders were also presented at

the beginning of each trial in order to ensure goal maintenance. The cute or threatening target stimulus was presented in a nine frame RSVP stream consisting of eight neutral animal stimuli which were randomly selected from the total pool of neutral stimuli. Each stimulus frame was presented for 100ms with no inter-stimulus interval. The target stimulus appeared at positions five, six, seven, or eight in the RSVP stream an equal number of times within each block, and was counterbalanced across conditions. The peripheral distractor stimulus was consistently presented two slides prior to the target at lag 2 on every trial.

The peripheral distractors were two images presented above and below the central stimulus position. One of these stimuli was always a neutral animal stimulus which was randomly selected from the pool of neutral animal images. The other distractor stimulus could either be a threatening animal, cute animal, or another neutral animal. Within each condition the distractor image appeared an equal number of times above and below the central stream. At the end of each trial, the participant typed out the animal they identified as the cute or threatening target using the keyboard and pressed 'Enter' to proceed to the next trial. The dependent variable was the percentage of trials that participants accurately reported the cute or threatening animal which had been presented.

Before the main task, participants completed a single eight trial practice block with four cute targets and four threat targets (the specific images used in these practice trials were different from the set used in the main experiment). For the main task, participants completed six blocks of 36 trials each, with a period of rest every two blocks, the duration of which was determined by the participant. The search condition blocks were presented in an alternating format (e.g. cute-threat-cute-threat-cute-threat). The block order was counterbalanced between participants, with half the participants completing a threat search block first. When blocks were not separated by a rest period, a text warning was presented for 3000ms alerting the participant that the search goal had changed. Other than search goal, which was

manipulated between blocks, all within participant factors were fully counterbalanced within each block. After completion of the study, participants completed self-report measures related to anxiety for exploratory purposes. However, given the sample sizes in relation to individual difference effects, this data is not reported here.

-----  
Insert Figure 1 about here  
-----

**Scoring.** The percentage of correctly identified animals was recorded as the outcome measure for analysis. In order to objectively score this measure, an excel formula was applied which marked a trial as correct when the spelling of the target animal matched the spelling of the response. To make sure that the responses were readable for this formula they were coded prior to the analysis, the coding rules and criteria can be found in Supplementary Materials 1. During the coding process the experimenter was blind to both the distractor conditions and the correct answers.

**Analytic strategy.** Data from Experiments 1, 2, 3b and 5 were significantly skewed (skewness ratio > 1.96) therefore an arcsine transformation was applied to the data. All statistics were performed upon the arcsine transformed data. For ease of interpretation, graphs are presented with untransformed data. We note the results remained unchanged with respect to patterns and significance when untransformed data were analysed. Analyses were performed using SPSS, and R-studio for Bayesian analyses (IBM Corp, 2016; R-studio team, 2015).

For all pair-wise comparisons between experimental conditions, 95% confidence intervals were bootstrapped (1000 samples), alongside conventional p-values (Field, 2013; Cumming, 2013). Hedges' g effect size was also calculated as a standardised effect size for all pairwise comparisons.

To supplement our main analysis we computed Bayes factors in order to determine whether any null effects were due to insensitivity or a true null effect. Further details of this analysis are reported in Supplementary Materials 2. A Bayes Factor compares evidence for the experimental hypothesis (threat relevant stimuli will result in greater attentional capture) and the null hypothesis (threat relevant stimuli will not result in attentional capture). Bayes factors ranges from 0 to infinity, values less than 1 indicate that there is support for the null hypothesis, whilst values of greater than 1 indicate that there is support for the experimental hypothesis. The strength of this evidence is indicated by the magnitude of the Bayes factor; values greater than 3 or less than .33 indicate substantial evidence for either the experimental or null hypothesis. A value closer to 1 suggests that any non-significant result is due to insensitivity and any difference is ‘anecdotal’ (Jeffrey, 1961; Dienes, 2008; 2011; 2014; 2016). All direct comparisons between conditions were tested using Bayes factors, however, p-values were also computed using two-way paired samples t-tests to facilitate comparison to previous results.

## Results and Discussion

Mean accuracy in each condition of search goal and distractor category can be seen in Table 1.<sup>1</sup> A 2×3 ANOVA with the factors of current search goal (cute/ threatening animal) and distractor category (cute/ threatening/ neutral animal) were performed on mean accuracy. This revealed that there was no significant difference in the accuracy with which participants

---

<sup>1</sup> Note that it is difficult to give a precise estimate of chance level performance in the identification task as this would depend on how many different animal identities participants were guessing from. If we were to assume that participants were guessing from the six target animals used in each search condition, chance performance would be only 16.67%. Given that participants had no prior knowledge of what the six animals would be, chance performance would likely be lower than this. Our finding of ~50% accuracy in the identification experiments therefore reflects responding well above chance.

identified cute versus threatening targets,  $F(1, 18) = 1.60, p = .222, \eta^2_p = .08$ . There was a significant main effect of distractor,  $F(2, 36) = 7.51, p = .002, \eta^2_p = .30$ , with the cute and threatening distractors resulting in lower performance than neutral distractors. Importantly, and consistent with the predicted goal-driven capture effect, this effect was qualified by a highly significant interaction between target and distractor,  $F(1.69, 30.47) = 16.11, p < .001, \eta^2_p = .48$  (Huynh-Feldt corrected).

In order to plot the effects more clearly, we created an affective distractor effect score by subtracting the accuracy when the distractor was cute or threatening from the neutral distractor condition, both for cute and threat search conditions (see Figure 2a). Performance when the distractor was a cute animal was lower when the target was also a cute animal, and a similar pattern was also observed for threatening animal distractors when the target was also a threatening animal. Thus, participants were significantly poorer at identifying the target when the distractor category matched the current search goal, as demonstrated by significantly greater distractor effects (i.e. the difference between neutral and affective distractors) when the affective distractor was goal-congruent versus incongruent. This was true for the threatening animal distractor effect,  $M = .70, SD = 9.07$  vs  $M = 9.87, SD = 11.11, t(18) = 2.60, p = .018, 95\% CI[2.54, 16.12], B_{H[0,15]} = 8.06$ , as well as the cute animal distractor effect,  $M = .39, SD = 10.67$  vs  $M = 12.75, SD = 11.17, t(18) = 4.08, p = .001, 95\% CI[6.34, 18.06], B_{H[0,15]} = 486.47$ .

The majority of errors consisted of naming an animal that was neither a distractor nor target, but similarly to the prior study by Wyble and colleagues (2013), on a percentage of trials the goal-congruent distractor was named in lieu of the target (15.54% in the threat search condition; 21.84% in the cute search condition; see Supplementary Materials 3 for a full break down of error type by condition in Experiments 1-3b).

---



-----  
Insert Table 1 about here  
-----

To explore evidence for both goal-congruent and goal-incongruent capture by threatening and cute animals in detail, we compared target identification accuracy between affective and neutral distractors within each search condition (see Table 1). As can be seen in Figure 2a, affective distractor effects, computed from affective versus neutral distractors, were only observed when the distractors were goal-congruent. Strikingly, there was no reduction in performance when the cute and threatening animal distractors were incongruent with the current search goal. The Bayes factors for both cute and threatening animal distractor effects are under .33 and hence confirm that the null results reflect an absence of attentional capture rather than insensitivity. Therefore, there was substantial evidence that, within our task, salient affective distractors only captured attention when they were congruent with current top-down search goals.

Note that although the affective ratings for the neutral animals were closer to those of the cute versus threat animals, this cannot explain our results. In the goal-congruent conditions, both cute and threat distractors produced robust interference effects of similar magnitude relative to neutral distractors. By contrast, neither produced interference in the goal-incongruent condition.

-----  
Insert Figure 2a and 2b about here  
-----

The results of Experiment 1 provide direct evidence that involuntary attentional capture by affective stimuli, both threatening and positive, can be induced via the adoption of a specific top-down goal, even when they appeared in completely task-irrelevant locations.

## **Experiment 2**

Experiment 2 sought to extend the findings of Experiment 1, by testing whether goal-driven capture by threat would generalise beyond the specific stimulus category (e.g. ‘threatening animals’) to the broader affective category (i.e. any form of threat). To test this possibility, we presented another widely used category of threat relevant stimuli as distractors, these being emotional faces. If the emotional faces captured attention more when they were congruent with the general affective content of the search goal (i.e. threatening animal search goal - fearful face distractor), this would imply the ability to adopt a broad attentional setting for an entire affective category which generalised automatically across conceptual boundaries.

We chose to present emotional faces as stimuli due to their universal recognition across individuals (Izard, 1994; Kohler et al., 2004). In regard to threat processing, specifically, we selected fearful faces because attentional biases to fear emerge in infancy, suggesting a rapidly learnt or innate signal of threat (Peltola, Hietanen, Forssman & Leppanen, 2013). Further, they have been found to reliably activate regions associated with automatic threat processing (i.e. the amygdala; Bishop, 2008), with this activation occurring more strongly than for other negative emotions such as anger (Fitzgerald, Angstadt, Jelsone, Nathan & Luan Phan, 2005). Thus, fearful faces are an ideal threat signal due to their universality and their strong relation to automatic threat processing and attention.

## **Methods**

**Participants.** Twenty participants were initially recruited, though 2 participants were excluded prior to analyses for taking an excessively long time to complete the search task

(over 50 minutes, compared to the typical task duration of 20-25 minutes).<sup>2</sup> The final sample consisted of 12 females and 6 males (Age:  $M = 21.78$ ,  $SD = 2.39$ ). The sample size was retained from our previous experiment. A power analysis using G\*power software indicated that this sample size would afford sensitivity to detect effect sizes above  $dz = .70$  with power of  $\beta = .80$  and an  $\alpha$  of .05 (Faul et al., 2007). Participants were recruited through the University of Sussex subject pool via an online advert. They were remunerated with course credits or a small cash payment.

**Stimuli and Procedure.** Stimuli were the same as those used in Experiment 1, with the exception that face stimuli were selected as distractors rather than animals. Twelve fearful faces, twelve happy faces, and twelve neutral faces were selected; they all shared the same 12 identities so were matched on every feature except emotion (Tottenham et al., 2009). As in previous instigations which found attentional biases towards fearful faces (e.g. Hodsoll, Viding & Lavie., 2011), we ovalled the faces to remove any non-emotional identifying features of the outline, such as hair style. To fill the opposite distractor location not occupied with the face distractor we presented one of twelve different skin patches created from close-ups of just the skin from the exemplars.

Due to the face stimuli being taller than animal images, distractors were presented to the left and right of the target in an upright position. In order to compensate for the increased distance from the centre of attention, the images were enlarged so they measured  $11.33^\circ \times 7.49^\circ$ . They were presented with a gap of  $.5^\circ$  between them and the central RSVP stream.

## Results and Discussion

---

<sup>2</sup> Including these two participants did not alter the significance or pattern of our findings.

The same 2×3 ANOVA was conducted as in Experiment 1, though the distractor conditions were now emotional faces (happy/ fearful/ neutral faces). In contrast to the results of Experiment 1, the analysis revealed no significant effects across Experiment 2, all  $p$ 's > .174,  $\eta^2_p < .10$ . Pairwise comparisons revealed that there was no evidence of generalisation of goal-driven capture across similar affective categories, even when the distractors were congruent with the search goal's general affective category, as can be seen in Figure 2b (see Table 1 for analyses). The Bayes factors all favoured the null but were nearer 1, therefore, the data were insensitive and required further evidence to draw a strong conclusion.

It, therefore, appears that in both Experiments 1 and 2 there was an unexpected absence of any attentional capture effects from either positive or threatening stimuli when these did not share the same specific affective category as the current search goal. It should be noted, however, that the distractors in Experiment 2 were presented further away from fixation than those in Experiment 1 to accommodate the stimulus dimensions. In order to allow a more direct comparison of the two distractor categories used in Experiments 1 and 2, further experiments were conducted in which both faces and animal distractors were presented in identical locations.

### **Experiment 3a and 3b**

The aim of Experiment 3a and 3b was to (1) replicate Experiment 1's finding of goal-driven attentional capture by affective stimuli, and (2) further test the possibility that this goal-driven attentional capture might generalise beyond the specific stimulus category (e.g. 'threatening animals') to the broader affective category (e.g. 'threat'), after controlling for distractor location. To allow direct comparison of these potential specific and more generalised goal-driven attentional capture effects, we incorporated both distractor categories into our task and presented both in the same parafoveal locations in Experiment 3a, and foveal locations in Experiment 3b. Participants performed the same central animal search task

as in Experiments 1 and 2, while ignoring distractors that were either threatening animals, fearful faces, or neutral animals and faces. Positive distractors were removed in order to focus specifically on the effect of different threat distractors on involuntary attention, which was the central aim of the current investigation. We expected to replicate Experiment 1's finding that threatening animal distractors would interfere with target identification only in the threatening animal search condition. It was unknown whether, having controlled for differences in distractor location, these contingent capture effects would now also generalise to the fearful faces (i.e. revealing interference from these affectively congruent stimuli exclusively in the threat search condition).

In Experiment 3b we presented the distractors in the central RSVP stream, rather than in peripheral or parafoveal locations, where generalisation may be more likely to occur due to greater visual processing at central target locations (Beck & Lavie, 2005). Additionally, all previous demonstrations of the EIB presented threatening distractors in a target location (e.g. Most et al., 2005). Presenting the distractors in the central stream would allow for a closer comparison to previous investigations which have found attentional capture by threat in an RSVP stream paradigm.

## Methods

### Participants.

**Experiment 3a.** Twenty participants were initially recruited for Experiment 3a, though one participant was excluded prior to analysis for accuracy being 3 SDs below the group mean, and another because of a programming error (12 females, 6 males; Age:  $M = 20.89$ ,  $SD = 2.65$ ).

**Experiment 3b.** Nineteen participants were initially recruited, though one participant was excluded prior to analysis for accuracy being 3 SDs below the group mean (16 females, 2 males; Age:  $M = 22.44$ ,  $SD = 4.83$ ).

The sample size for both experiments was retained from our previous experiments, and afforded the sensitivity to detect effects above  $dz = .70$  ( $\beta = .80$ ;  $\alpha = .05$ ; Faul et al., 2007). Participants were recruited through the University of Sussex subject pool via an online advert. They were remunerated with course credits or a small cash payment.

### **Stimuli and Procedure.**

**Experiment 3a.** The stimuli and procedure in Experiment 3a were identical to Experiments 1 and 2, though in order to compare the effect of emotional faces and threatening animals within a single experiment, the following changes were made to the design: A  $2 \times 2 \times 2$  within-subjects design was used: Target type (cute/ threat animal)  $\times$  Distractor type (animal/ face)  $\times$  Distractor valence (threat related/ neutral). Additionally, all images were reduced in size in order to place them in parafoveal vision ( $>2.5^\circ$  eccentricity), rather than peripheral vision ( $> 5^\circ$ ; cf. Toet & Levi, 1992). This meant that images in the central RSVP stream measured  $3.44^\circ \times 2.29^\circ$ , and distractors measured  $2.98^\circ \times 4.58^\circ$  visual angle at 59cm viewing distance from the screen. The distractors were presented to the left and right of the central RSVP stream with a gap of  $.5^\circ$  between the central image and the distractor. The order of distractors and targets was pseudo randomly generated in order to prevent the distractor being the same animal as the target, or regular pairings of distractor and target emerging by chance.

Stimuli were taken from the images used in Experiments 1 and 2. The neutral animal distractors were six images of six different animals (capybara, sheep, pig, catfish, goose, pigeon), these exemplars never appeared as part of the central stream. Similarly, six separate threatening animals were selected from those used in Experiment 1. Six fear and six neutral faces were selected to be distractors from those used in Experiment 2. Both fear and neutral faces shared the same individual identities, meaning that the only difference was their emotion. As in Experiments 1 and 2 one distractor image appeared per trial - the opposite

side distractor location was occupied with an oval patch of skin or animal texture (e.g. close-up of fur or feathers). Twelve skin and twelve animal texture exemplars were created from close up images of faces and animals sourced from Google images. Texture patches were presented only alongside their congruent distractor type (i.e. skin patch alongside face distractor), and were randomly selected across the block. To remove shape differences between the animal and face distractors, all distractors were ovalled leaving only the key features of both animals and faces. They were both presented in an upright position during the experiment.

Six threatening animal images and six cute animal images were selected to be targets from those used in Experiments 1 and 2; each target category was made up of the same six different animals presented.<sup>3</sup> Neutral filler animals were made up of 192 images selected to appear in the central RSVP. Participants completed four blocks of 48 trials each with the cute search blocks and threat search blocks structured in an alternating format (i.e. cute-threat-cute-threat), the order of which was counterbalanced between participants. Within both cute and threat search blocks the four types of distractor were presented with equal probability, these appeared equally to the left and the right of the target.

After the RSVP task, participants rated all target and distractor images, in a random order, along dimensions of arousal and valence using a self-assessment manikin (see Table 2; Bradley & Lang, 1994). All participants from Experiments 3a, 3b, 4 and 5 ( $N = 79$ ) completed the rating task, which was programmed in Inquisit 5 software (Millisecond, 2016). Ratings from each individual experiment produced a similar pattern of results.

---

<sup>3</sup> The crocodile stimuli were replaced due to very poor performance in identifying these targets in Experiment 1 and 2. These were replaced with images of crocodiles which were more visible.

**Experiment 3b.** The task and procedure were nearly identical to Experiment 3a with the exception that the distractor appeared in the central stream. These distractors were marked as task-irrelevant by presenting them as a  $1.53^\circ \times 2.29^\circ$  oval, which was presented within a grey rectangle amongst the other stimuli which were all complete rectangular images. This change resulted in one fewer neutral filler image per trial, leaving a total of 168 neutral animals images selected to appear across the experiment. Additionally, for the purposes of counterbalancing, the number of target locations in the RSVP stream was reduced to positions six, seven and eight.

-----

Insert Table 2 about here

-----

-----

Insert Table 3 about here

-----

## Results and Discussion

For both Experiments 3a and 3b, identification accuracy across the eight conditions (see Table 3) was analysed in a  $2 \times 2 \times 2$  repeated measures ANOVA: search goal (cute/ threatening animal)  $\times$  distractor type (face/ animal)  $\times$  distractor valence (threat related/ neutral). In order to clearly illustrate the results, we plotted the affective distractor effects for both Experiment 3a and 3b (see Figures 3a and 3b); to compute these effects, we only subtracted the affective distractor accuracy from the neutral distractor accuracy of the same type.

The analysis for Experiment 3a revealed that there was a non-significant difference between accuracy in the cute animal search goal or threatening animal search goal,  $F(1, 17) = 1.37, p = .259, \eta_p^2 = .07$ , but the same difference was significant for Experiment 3b,  $F(1, 17)$



$= 14.88, p = .001, \eta^2_p = .47$ . There was a non-significant difference between distractor types for Experiment 3a,  $F(1, 17) = .42, p = .524, \eta^2_p = .02$ , however there was a significant difference between faces and animal distractors within Experiment 3b,  $F(1, 17) = 12.03, p = .003, \eta^2_p = .40$ . The effect of distractor valence was marginally significant for Experiment 3a,  $F(1, 17) = 4.24, p = .055, \eta^2_p = .20$ , and significant in Experiment 3b,  $F(1, 17) = 22.57, p < .001, \eta^2_p = .57$ .

For Experiment 3a, search goal did not significantly interact with distractor type,  $F(1, 17) = .69, p = .417, \eta^2_p = .04$ , though this effect was marginal for Experiment 3b,  $F(1, 17) = 3.89, p = .065, \eta^2_p = .19$ . For both experiments, search goal significantly interacted with distractor valence, 3a:  $F(1, 17) = 7.14, p = .016, \eta^2_p = .30$ ; 3b:  $F(1, 17) = 28.78, p < .001, \eta^2_p = .63$ . Distractor type also interacted with distractor valence in both experiments, 3a  $F(1, 17) = 6.36, p = .022, \eta^2_p = .27$ ; 3b:  $F(1, 17) = 26.26, p < .001, \eta^2_p = .61$ . Critically, for both Experiments 3a and 3b, the three-way interaction was significant, 3a:  $F(1, 17) = 7.25, p = .015, \eta^2_p = .30$ ; 3b:  $F(1, 17) = 29.13, p < .001, \eta^2_p = .63$ .

In both experiments the threatening animal distractor effects were greater when they were congruent with the specific search goal, compared to when they were incongruent with the search goal, Experiment 3a:  $M = -2.31, SD = 7.73$  vs  $M = 11.57, SD = 8.99, t(17) = 4.70, p < .001, 95\% CI[-19.52, -8.10], B_{H[0,13]} = 4305.04$ ; Experiment 3b:  $M = .05, SD = 12.56$  vs  $M = 31.23, SD = 16.45, t(17) = 6.85, p < .001, 95\% CI[-39.08, -22.15], B_{H[0,13]} = 45741033$ . Thus, the specific interaction between search goal and the threatening animal distractor effects found in Experiment 1 was replicated in both experiments

-----  
Insert Figure 3a and 3b about here  
-----

Interestingly, the magnitude of the distractor effect for the goal-congruent threatening animals in Experiment 3b was significantly larger than that found in Experiment 3a (corrections were made for lack of equal variances between groups),  $t(26.32) = 4.45$ ,  $p < .001$ ,  $g = 1.32$ , 95% CI [10.97, 28.02]; all other  $p$ 's  $> .501$ . This supports the idea that the distractors were processed more deeply at target locations, and is in keeping with previous evidence that centrally presented distractors interfere more with task performance than more peripheral distractors (Beck & Lavie, 2005).

Follow-up pairwise comparisons (see Table 3) revealed that, in both experiments, threatening animal distractors resulted in significantly lower performance relative to neutral animals in the threatening animal search condition, but there was no evidence of attentional capture in the cute search condition. There was, also, no difference between fearful face and neutral face distractors in either the threatening animal search condition or the cute animal search condition. All null effects produced Bayes Factors which were below .33, thus revealing that there was substantial evidence in favour of the null hypothesis when threat distractors were incongruent with the current search goal. Such a null finding was unexpected, especially for Experiment 3b, given that previous investigations have demonstrated goal-incongruent capture by threat in a very similar task (e.g. Most et al., 2005) – we discuss the possible reasons for this difference in the General Discussion.

### **Experiment 4**

Experiments 2-3b found no evidence of attentional capture from threat related faces, despite using the exact same face stimulus set that has previously elicited involuntary attentional capture in a multitude of experimental tasks (Tottenham et al., 2009). However, in our current experiments, the faces never directly matched the task's top-down goal. Experiment 4 sought to test whether the face stimuli would be capable of capturing attention in the current paradigm when they match top-down task goals. To this end we modified our

task search goals, so that instead of searching the central stream for cute or threatening animals, participants were now instructed to search for happy or fearful emotional faces. We presented identical distractor stimuli to those used in Experiments 3a and 3b, in order to see whether current search goals could now induce involuntary attentional capture by emotional faces.

## Methods

**Participants.** A sample of 18 participants were recruited for this experiment (11 females, 7 males; Age:  $M = 21.06$ ,  $SD = .54$ ). The sample size was retained from our previous experiments, and afforded the sensitivity to detect effects above  $d_z = .70$  ( $\beta = .80$ ;  $\alpha = .05$ ; Faul et al., 2007). Participants were recruited through the University of Sussex subject pool via an online advert, and were remunerated with course credits or a small cash payment.

**Stimuli and Procedure.** The experimental design and structure was identical to Experiment 3a, with the exception of the following changes. Firstly, participants were instructed to search for happy faces, instead of cute animals, and fearful faces instead of threatening animals. There were two blocks of 96 trials, one for the happy face search, and one for the fearful face search. The order of these blocks was counterbalanced between participants. An additional change to the paradigm was that participants had to identify whether the emotional face was present or absent on each trial. They responded using the ‘c’ and ‘m’ keys, the key-response assignment was counterbalanced between participants. The target was present on half of trials, when it was absent, the target was replaced by an upright neutral face. Pilot testing revealed that participants were performing at ceiling, thus the stimulus presentation time was reduced to 83ms per frame with no inter-stimulus interval. Unlike the RSVP of previous experiments which were composed entirely of animal images, the neutral filler stimuli were composed of two neutral animals selected from the previous pool of neutral images, three inverted faces, and either three or four upright faces, depending

on whether the target was present or absent. The multiple types of filler stimuli were presented in a random order in each trial, their purpose was to increase the difficulty of the task. In total, 48 neutral animals were presented in the central RSVP stream, alongside 72 upright faces and 72 inverted faces all with different identities. The neutral face stimuli were taken from the Productive Aging Laboratory Face database (Minear & Park, 2004). The target stimuli consisted of three happy faces and three fearful faces of the same identities. These were taken from the NimStim database (Tottenham et al., 2009). Distractor and target faces were selected so that different ethnicities and genders appeared equally across distractors and targets in each individual condition. Additionally, neutral filler images were selected so that male and female faces were equally represented, and that different ethnicities were presented approximately equally. Thus, the face stimuli appeared as a heterogeneous stream of facial features. An eight-trial practice block preceded the task with equal happy and fearful targets and equal present and absent trials. Stimuli presented in the practice were not presented in the rest of the experiment, and distractors in the practice block consisted of black ovals.

## Results and Discussion

Unlike Experiments 1-3b, the signal detection measure  $A'$  was the dependent variable used to measure target detection rather than % accuracy. Signal detection measures can control for biases introduced with binary responses and are often used in RSVP designs (e.g. Brown, Duka & Forster, 2018; Failing & Theeuwes, 2015).  $A'$  was computed based on the proportion of hits (i.e. correctly responding 'present' when the target was present) and false alarms (i.e. incorrectly responding 'present' when target was absent) made during the present/absent task response (see Table 4; Stanislaw & Todoroff 1999; Zhang & Mueller 2005).  $A'$  ranges from .5, which indicates that a signal cannot be distinguished from noise (i.e. chance detection), to 1, which corresponds to perfect detection of the target.  $A'$

data from Experiment 4 was significantly skewed, however, due to arcsine transformations actually increasing skewness we used the original data. We note that the bootstrapped 95% confidence intervals, which are robust to violations of normality, corroborated the conclusions based on the  $p$ -values calculated using skewed data (see Table 4; Field, 2013).

As in Experiment 3a and 3b, we conducted a  $2 \times 2 \times 2$  ANOVA, although the search goal factor was changed to happy/fearful search conditions. The main effect of search goal was non-significant,  $F(1, 17) = .06, p = .809, \eta^2_p < .01$ . The main effect of distractor type was, however, significant,  $F(1, 17) = 8.07, p = .011, \eta^2_p = .32$ ; and, in a striking reversal of previous results, the face distractors resulted in lower performance overall compared to animal distractors. The main effect of distractor valence was also significant,  $F(1, 17) = 15.78, p = .001, \eta^2_p = .48$ , with threat related (i.e. fearful face or threatening animal) distractors resulting in lower performance than neutral distractors. Critically, the distractor type interacted with distractor valence,  $F(1, 17) = 5.86, p = .027, \eta^2_p = .26$ , revealing that interference was observed from fearful versus neutral faces but no difference between threatening animals versus neutral animals. The interactions between current search goal and distractor type, and search goal and distractor valence failed to reach significance, as did the three-way interaction between these variables, all  $p$ 's  $> .175, \eta^2_p < .11$ .

-----  
Insert Table 4 about here  
-----  
-----

-----  
Insert Figure 4 about here  
-----

As can be seen clearly in Table 4, Bayesian pairwise comparisons revealed that the absence of the three-way interaction was due to fearful faces resulting in lower detection

sensitivity relative to neutral faces across both the emotional face search conditions. The difference between the distractor effects between the search goal condition was non-significant,  $M = .05$ ,  $SD = .10$  vs  $M = .08$ ,  $SD = .09$ ,  $t(17) = 1.29$ ,  $p = .213$ , 95% CI $[-.08, .01]$ ,  $B_{H[0,.11]} = .79$ , though Bayes factor showed that the data were insensitive and the magnitude of the distractor effects was in the expected direction. In contrast to all previous experiments, the threatening animal distractor effects did not significantly differ depending on the emotional face search goals,  $M = .01$ ,  $SD = .05$  vs  $M = .02$ ,  $SD = .07$ ,  $t(17) = .72$ ,  $p = .480$ , 95% CI $[-.06, .2]$ ,  $B_{H[0,.11]} = .38$ .

The unexpected significant attentional capture by fearful faces in the happy search condition might at first glance be assumed to be evidence of stimulus-driven attentional capture. However, it is hard to reconcile a stimulus-driven interpretation of this effect with the fact that, across four experiments, we only observed this effect when the task search categories were changed from animals to faces. This dependence of the attentional capture effect on the central task stimulus category points to a goal-driven rather than stimulus-driven mechanism.

Why then did we not find a significant within-subjects goal-driven effects on attentional capture in Experiment 4? We speculate that this may be due to the increased overlap in visual features between the two face affective categories. Unlike the visually distinct cute and threatening animal categories used in previous experiments, happy and fearful faces share common features such as visible teeth. It has been found that, when possible, participants search for a single salient visual feature of an emotional face, especially in a perceptually demanding task (Calvo, Fernandez-Martin & Nummenmaa, 2012; Horstmann, Lipp & Becker, 2012). Indeed, removing these key salient features, such as teeth or eyes, from emotional faces reduces the efficient detection of these faces when participants are searching for them (Lee, Susskind & Anderson, 2013). Participants would only have to

hold a top-down search goal for salient mouths or eyes amongst the neutral faces to complete the task. This, therefore, would lead to goal-driven capture by all emotional faces possessing these shared features.

### Experiment 5

Out of five experiments, Experiment 4 was the first to use a present/absent response format. Before comparing across our experiments, it was important to check whether our key finding of goal-driven attention capture (as observed in all three prior experiments using animal distractors) could be replicated using this response format. We therefore conducted a nearly identical replication of Experiment 4 with the change that the central search task used animal stimuli rather than faces. We preregistered this replication on the Open Science Framework (OSF: [osf.io/fkexj](https://osf.io/fkexj)); any deviations from the preregistered procedure and analysis are noted below. The key prediction was that searching for threatening animals would result in a greater distractor effect (neutral minus threatening animal distraction) compared to when participants were searching for cute animals.

### Methods

**Participants.** Initially 26 participants were recruited, though one was excluded for scoring 50% or below on detecting one of the target categories. Thus, 25 participants were carried forward to analysis, these included 16 female and 9 male participants, the mean age of which was 25,  $SD = 4.07$ . Sample size was calculated using an a priori power analysis in G\*power using the smallest effect size from the key interaction terms taken from across Experiments 1, 3a and 3b (Faul et al., 2007). This interaction effect size was calculated by comparing the distractor effect of the goal-incongruent threatening animals versus the distractor effect of the goal-congruent threatening animals. The largest sample predicted was 24 which was found using the interaction term from Experiment 1 for threatening animals,  $d_z$

$= .60$ ,  $\beta = .80$ ,  $\alpha = .05$ , two more participants were recruited (following our pre-registered plan) to pre-empt potential exclusions. Participants were recruited through the Birkbeck University of London subject pool via an online advert, and were remunerated with a small cash payment.

**Stimuli and Procedure.** The experiment was identical to Experiment 4 with the exception of the following changes. The stimuli were presented on a 24inch Dell monitor with the resolution set to  $1920 \times 1080$  with a refresh rate of 60Hz. The targets and distractor stimuli were identical to those used in Experiment 3a; each image was presented for 100ms. The participants had to search for cute animals and threat animals in four separate blocks of 96 trials (e.g. cute – threatening – cute – threatening). The order of these blocks was counterbalanced between participants. The neutral animal filler stimuli which made up the central RSVP stream consisted of 204 neutral animals taken from Experiment 1.

## Results and Discussion

Though we pre-registered that both % accuracy and  $A'$  would be the dependent variables we only report the results with  $A'$  in the main paper in order to account for potential response bias with the binary response. The analyses conducted with accuracy as the dependent variable does not alter the significance or pattern of results and is reported in Supplementary Materials 4. The same  $2 \times 2 \times 2$  ANOVA from Experiment 4 was performed on the arcsine transformed  $A'$  data (see Table 4); this revealed that there was a significant effect of search goal,  $F(1,24) = 4.87$ ,  $p = .035$ ,  $\eta_p^2 = .17$ , with threatening animals detected more accurately than cute animals. There was also a significant difference between distractor type,  $F(1,24) = 8.72$ ,  $p = .007$ ,  $\eta_p^2 = .27$ , with animal distractors resulting in lower accuracy. The main effect of distractor valence was, however, non-significant,  $F(1,24) = .67$ ,  $p = .665$ ,  $\eta_p^2 = .01$ .



The interaction between search goal and distractor type was non-significant,  $F(1,24) = 1.83, p = .189, \eta_p^2 = .07$ . However, the interaction between search goal and distractor valence was approaching significance,  $F(1,24) = 4.21, p = .051, \eta_p^2 = .15$ , reflecting reduced  $A'$  when threatening stimuli were congruent with the current search goal. There was no significant interaction between distractor type and valence,  $F(1,24) = .11, p = .739, \eta_p^2 = .01$ . The three-way interaction also failed to reach significance,  $F(1,24) = 3.13, p = .090, \eta_p^2 = .12$ , although as seen in Figure 5 a trend was observed in line with that seen in Experiments 3a and 3b.

Critically, the specific pairwise comparison which the experiment was powered to detect (i.e. the difference between the goal-incongruent threat distractor effect and goal-congruent threat distractor effect) was significant,  $M = -.02, SD = .07$  versus  $M = .04, SD = .08, t(24) = 2.63, p = .015, 95\% CI[-.02, .11], B_{H[0,.11]} = 7.99$ . Thus, the a priori comparison providing the key test of our hypothesis confirmed that there was an increase in attentional capture by threatening animals when congruent with the current search goals (see Figure 5).

-----  
 Insert Figure 5 about here  
 -----

Furthermore, as predicted, the preregistered pairwise comparisons between the threat related distractors and the matched neutral distractor revealed that the only condition which showed a difference was when the threatening animal distractor was congruent with the search goal. This same difference between the animal distractors was non-significant in the cute search condition. The fearful face distractors were non-significantly different from the neutral face distractors in both the threatening animal search conditions. Bayes factors for all non-significant comparisons strongly favoured the null hypothesis, whilst only the goal-congruent condition showed evidence for the experimental hypothesis, as we predicted in our preregistration.

Hence the patterns of significant attentional capture (i.e. distractor interference) observed in Experiment 5 were consistent with the findings of our four other experiments involving animal search: Animal distractors captured attention only when goal-congruent, and face distractors did not capture attention at all. The discrepant findings from Experiment 4 are thus unlikely to reflect the change in response mode. Rather, the only difference between Experiment 4 and 5 was the change from animal search to face search. As such, this change is implicated as the critical factor determining why face distractors interfered in Experiment 4 but not in the other experiments. Indeed, exploratory comparisons between experiments revealed that the average fearful face distractor effect was significantly larger during face search (Experiment 4,  $M = .07$ ,  $SD = .08$ ) compared to the average fear distractor effect during animal search conditions (Experiment 5,  $M < .01$ ,  $SD = .08$ ),  $t(41) = 2.83$ ,  $p = .007$ , 95% CI[.02, .12].

### **Internal Meta-analysis**

To demonstrate the overall evidence for goal-driven (i.e. goal-congruent) attentional capture, above and beyond any stimulus-driven (i.e. goal-incongruent) capture, we conducted an internal meta-analysis across our six experiments. Full details of this meta-analysis are presented in Supplementary Materials 5.

We conducted a multivariate moderation meta-analysis of the difference between the average threat distractor accuracy and the average neutral distractor accuracy in each experiment, with goal-congruence as a moderating factor, and Experiment as a random effects factor. The analysis was computed using a restricted maximum likelihood estimate in the Metafor package for R (Viechtbauer, 2010). This analysis confirmed a highly significant effect of goal congruence on attentional capture by threat,  $g = .61$ ,  $p = .003$ , 95% CI [.20, 1.02], with goal-congruent threat distractors resulting in a greater decrement in performance versus neutral distractors, compared to when the same distractors were goal-incongruent.

We then investigated the magnitude of both the goal-congruent and goal-incongruent effects independently. A DerSimonian-Laird random effects model (DerSimonian-Laird, 1986) revealed a large and significant goal-driven effect with strong evidence favouring the experimental hypothesis, ( $k = 5$ ,  $N = 98$ ),  $g = .94$ ,  $p = .020$ , 95% CI [.15, 1.74],  $B_{H[0,15]} = 4458246 \times 10^7$ , and a far weaker non-significant difference between the average of all neutral distractors and threat distractors when they were goal-incongruent, ( $k = 6$ ,  $N = 116$ ),  $g = .17$ ,  $p = .322$ , 95% CI [-.17, .52],  $B_{H[0,15]} = .54$ . In Experiment 4, we have discussed the possibility that participants had a general goal for emotional faces, thus capture may not be entirely goal-incongruent. Calculating the meta-analysis without Experiment 4 led to the goal-incongruent effect size dropping to near zero and Bayes factor signifying a sensitive null finding, ( $k = 5$ ,  $N = 98$ ),  $g = .01$ ,  $p = .960$ , 95% CI [-.27, .29],  $B_{H[0,15]} = .13$ .

### General Discussion

The present study reveals that involuntary attentional capture by threatening stimuli can be induced by manipulating current top-down search goals. Across six experiments, involuntary attentional capture by threatening stimuli only occurred when they shared the same category as the participants' current search goal (e.g. threatening animals). This was also true for positive stimuli. This pattern was observed across the visual field, with peripheral, parafoveal, and centrally positioned distractors. These findings demonstrate that attentional capture by affective stimuli such as threat could plausibly be accommodated within the goal-driven/stimulus-driven theoretical dichotomy as an involuntary phenomenon driven by goal-driven attention.

An alternative recent approach to accommodating affective stimuli into prominent models of attention proposes a third mechanism based on low-level inter-trial 'selection history' (Awh et al., 2012; Theeuwes, 2013). In previous investigations, removing primed trials substantially reduced the 'goal-driven' distractor effect (Lamy & Kristjánsson, 2013).

However, analyses of five of our experiments, where we found evidence of goal-driven capture, revealed that the removal of trials where the distractor had been preceded by a visually similar target did not abolish the decrement in accuracy (all  $p$ 's < .047, Hedges'  $g$  > .32,  $B$ 's > 3.85; see Supplementary Materials 6 for further details of this analysis). Further, analysis of Experiments 4 and 5, where there were sufficient trials to compare primed versus unprimed distractors, revealed that there was no significant difference in accuracy (all  $p$ 's > .232, Hedges'  $g$  < .31,  $B$ 's < 1.04). Thus, an intertrial priming account does not appear to fit our data. This is not to say that selection history is not found in previous tasks, in which unfamiliar coloured shapes are used as stimuli in relatively brief trials (e.g. ~3s; Theeuwes, Reiman & Mortier, 2006). It may be, however, that the long trial duration (~5s) in our task allowed visual activation from priming to dissipate; or alternatively, the more complex images and scenes used in the current task are less susceptible to low-level feature priming. Nevertheless, our results support the possibility of a parsimonious account of involuntary attentional capture by affective stimuli, which positions it within the existing goal-driven/stimulus-driven framework.

Our findings extend recent work on task-relevance (e.g. Everaert et al., 2013; Stein et al., 2009; Vogt et al., 2013; Lichtenstein-Vidne et al., 2012; 2017; Vromen et al., 2016) by providing direct evidence for a causal role of top-down goals in involuntary capture by threat. Our findings demonstrate that goal-driven affective capture is spatially global, occurring even in entirely task-irrelevant locations. On the other hand, it appeared to be rather specific to the particular conceptual category of the search target: Searching for one type of threatening stimulus (threatening animals) did not induce capture by other forms of threat (fearful faces), and vice versa. This contrasts with previous studies that have found contingent capture from stimuli conceptually associated with a target category (e.g. other clothing images capturing attention during a search for jeans; Nako, Wu, Smith & Eimer, 2014). This might imply a

distinction between the effects of semantic versus affective categories on attention, and an important boundary effect for goal-driven threat capture. In other words, a general desire to detect threat might not be sufficient for goal-driven processes to induce threat capture.

It should be considered, however, whether the current task context encouraged participants to adopt a narrower search goal than was strictly necessary to perform the task (e.g. searching specifically for threatening animals rather than simply for threat content). Future research could clarify this issue by adapting our paradigm to include threatening targets from multiple categories.

### **The Role of Goals in Clinical Attentional Biases**

Viewing attentional capture by affective stimuli as a form of goal-driven attentional bias also has interesting implications for understanding the attentional biases seen in relation to anxiety and other clinical conditions. Such biases have traditionally been accounted for in terms of the clinical syndrome or behaviour increasing ‘bottom-up’ responsivity to certain stimuli (e.g. Bishop, 2007; 2008; Mogg & Bradley, 2016; Cisler & Koster, 2010). However, it seems plausible that some individuals, such as those who are highly anxious, would consider certain affectively relevant goals to be highly important, and hence be more likely to voluntarily adopt these goals (see Wells & Matthews, 1994).

Interestingly, the specificity of affective contingent capture in the present study is consistent with patterns observed in relation to attentional biases: A recent meta-analysis of the attentional bias to threat in anxiety disorders concluded that threatening stimuli were prioritised more when they were congruent with an individual’s specific anxiety disorders (e.g. angry face for social anxiety) compared to when they were incongruent (Pergamin-Hight, Naim, Bakermans-Kranenburg, van IJzendoorn & Bar-Haim, 2015). Hence, our proposed goal-driven account of attentional capture by threatening stimuli neatly accommodates established patterns of attentional biases in clinical samples, and has the

important implication that these could plausibly be driven by personal concerns and goals rather than unconditional biases dictated merely by affective associations.

We therefore suggest that a goal-driven mechanism could conceivably provide an alternative account of previous laboratory demonstrations of seemingly stimulus-driven attentional capture by affective stimuli, in terms of participants' personal long-term goals overriding instructed task goals. During laboratory demonstrations of attentional biases, threat stimuli are typically repeatedly presented in the context of an undemanding task, creating the possibility that anxious participants might begin to neglect task goals and 'look out' for threat.

It should be noted, however, that goal-driven attentional capture has thus far only been demonstrated when individuals are actively, and temporarily, maintaining a search goal in line with task instructions (cf. Soto, Hodsoll, Rotshtein & Humphreys, 2008). More work is required to confirm whether ongoing (but perhaps more passive) personal priorities can influence attentional settings and hence induce similar goal-driven capture effects. Interesting preliminary evidence for this comes from a study by Purkis, Lester and Field (2011), who found that biases akin to those found among spider phobics, in relation to images of spiders, were found among fans of the television show 'Dr Who', in relation to 'Dr Who' related images, despite searching an unrelated category of stimuli (i.e. horses). This finding is at odds with the traditional view of attentional biases for threat as being hard-wired and stimulus-driven, but is compatible with the view of these biases as reflecting an involuntary consequence of long-term goal-driven attentional settings.

### **Can Capture by Affective Stimuli Ever Be Purely Stimulus-Driven?**

We note that although we did not find evidence of stimulus-driven attentional capture across our experiments, our investigation was not designed to test for such effects, and our null results do not rule out the possibility that this might occur under some circumstances. It

may be that certain features of our paradigm may have reduced sensitivity to stimulus-driven effects. First, our use of affective stimuli as targets may have increased their ability to “survive” the attentional blink induced by distractors (cf. Reineke, Rinck & Becker, 2008; Raymond & O’Brien, 2009; Yokoyama, Padmala & Pessoa, 2015). Second, our task was rather perceptually demanding, raising the possibility that it might have filled perceptual capacity to the exclusion of distractor processing (e.g., Lavie, 2005, Forster & Lavie, 2008, Bishop, Jenkins & Lawrence, 2007). Third, our task used a clear predefined and specific goal, which in combination with expected distractors may have boosted the ability to override any stimulus-driven effects (Glickman & Lamy, 2018; Grimshaw et al, 2018).

We were surprised that we did not replicate the goal-incongruent capture by threat, especially when we moved the distractor into the same RSVP stream as the target, as has been done with previous research showing a goal-incongruent EIB in an RSVP task (Most et al., 2005; Smith et al., 2006; Kennedy & Most, 2012; Kennedy et al., 2018; Kennedy et al., 2014). We speculate that other features of the previous tasks’ design may account for the striking differences between our findings and the apparently stimulus-driven effects previously observed. Across all demonstrations of EIB, participants are required to search one or two RSVP streams for a neutral rotated image amongst upright images and identify which way it is rotated (left/right), whilst ignoring the threat distractor appearing prior to the target. Due to the target always being defined as the only rotated image in the RSVP stream, with no knowledge of the exact rotation prior to the target appearing, it is not possible to adopt a feature specific attentional setting and participants might instead rely on a perceptually undemanding odd-one-out search goal (see Bacon & Egeth, 1994). Both target uncertainty and low perceptual load are conditions in which fearful face stimuli capture attention independent of instructed task-goals, whilst providing participants with a predefined

constant target set or perceptually demanding tasks has been found to abolish threat capture effects (Glickman & Lamy, 2018; Bishop et al, 2007).

Future research should clarify whether stimulus-driven effects would be observed in a less perceptually demanding version of our task, when target features are emotionally neutral and inconsistent across trials. For now, we note that our current results clearly demonstrate that the goal-driven capture was substantial enough to override all of these factors. As such, although stimulus-driven capture may be possible, it may be confined to more limited situations than goal-driven capture.

It might alternatively be argued that ‘stimulus-driven’ effects are only found among certain individuals. Bar-Haim et al. (2007) found that in a meta-analysis of 172 studies, threat only reliably captured attention in anxious individuals. However, such biases are not necessarily indicative of stimulus-driven attentional capture – as discussed above, such biases could in fact reflect the participant’s own long-term concerns and goals resulting in the momentary prioritisation of these stimuli over the current task goals. Regardless of whether low-level stimulus features alone can elicit attentional capture by threat, it is important to note evidence that such features can moderate this phenomenon. Indeed, many affective categories are characterised by their low-level stimulus features. For instance, individuals intentionally tune attention towards salient facial features, like eyes and teeth, when instructed to search for emotional faces amongst other stimuli (Lee et al., 2013; Horstmann et al., 2012).

## **Conclusions**

To conclude, we have demonstrated that current search goals for affective categories, specifically threatening stimuli, can induce involuntary attentional capture by these stimuli. As such, our data provides direct experimental evidence for a goal-driven account of this form of involuntary attentional capture. Our findings have implications both for theoretical



models of attention, and for understanding the attentional biases seen in relation to anxiety disorders.

### References

- Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, 16, 437-443.
- Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., ... & Van Ijzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychological Bulletin*, 133, 1-24.
- Beck, D. M., & Lavie, N. (2005). Look here but ignore what you see: effects of distractors at fixation. *Journal of Experimental Psychology: Human Perception and Performance*, 31, 592-607.
- Bishop, S. J. (2007). Neurocognitive mechanisms of anxiety: an integrative account. *Trends in Cognitive Sciences*, 11, 307-316.
- Bishop, S.J. (2008). Neural mechanisms underlying selective attention to threat. *Annals of the New York Academy of Sciences*, 1129, 141-152.
- Bishop, S. J. (2009). Trait anxiety and impoverished prefrontal control of attention. *Nature Neuroscience*, 12, 92-98.
- Bishop, S. J., Jenkins, R., & Lawrence, A. D. (2007). Neural processing of fearful faces: effects of anxiety are gated by perceptual capacity limitations. *Cerebral Cortex*, 17, 1595-1603.
- Brown, C. R. H., Duka, T., & Forster, S. (2018). Attentional capture by alcohol-related stimuli may be activated involuntarily by top-down search goals. *Psychopharmacology*, 235, 2087–2099.
- Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25, 49-59.

- Buschman, T. J., & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science*, 315, 1860-1862.
- Calvo, M. G., Fernández-Martín, A., & Nummenmaa, L. (2012). Perceptual, categorical, and affective processing of ambiguous smiling facial expressions. *Cognition*, 125, 373-393.
- Carretié, L. (2014). Exogenous (automatic) attention to emotional stimuli: a review. *Cognitive, Affective, & Behavioral Neuroscience*, 14, 1228-1258.
- Cisler, J. M., & Koster, E. H. (2010). Mechanisms of attentional biases towards threat in anxiety disorders: An integrative review. *Clinical Psychology Review*, 30, 203-216.
- Cumming, G. (2013). The new statistics why and how. *Psychological Science*, 25(1), 7-29.
- DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7(3), 177-188.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193-222.
- Dienes, Z. (2008). Understanding psychology as a science: An introduction to scientific and statistical inference. Hampshire, England: Palgrave Macmillan. Retrieved from [http://www.lifesci.sussex.ac.uk/home/Zoltan\\_Dienes/inference/](http://www.lifesci.sussex.ac.uk/home/Zoltan_Dienes/inference/)
- Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on? *Perspectives on Psychological Science*, 6, 274-290.
- Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, 5, 781.
- Dienes, Z. (2016). How Bayes Factors change scientific practice. *Journal of Mathematical Psychology*, 72, 72-89.

- Everaert, T., Spruyt, A., & De Houwer, J. (2013). On the malleability of automatic attentional biases: Effects of feature-specific attention allocation. *Cognition & Emotion*, 27, 385-400.
- Everaert, T., Spruyt, A., Rossi, V., Pourtois, G., & De Houwer, J. (2014). Feature-specific attention allocation overrules the orienting response to emotional stimuli. *Social Cognitive and Affective Neuroscience*, 9, 1351-1359.
- Failing, M. F., & Theeuwes, J. (2015). Nonspatial attentional capture by previously rewarded scene semantics. *Visual Cognition*, 23(1-2), 82-104.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G\* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175-191.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics (4th ed.)*. London: Sage.
- Fitzgerald, D. A., Angstadt, M., Jelsone, L. M., Nathan, P. J., & Phan, K. L. (2006). Beyond threat: amygdala reactivity across multiple expressions of facial affect. *Neuroimage*, 30(4), 1441-1448.
- Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 1030-1044.
- Folk, C. L., Leber, A. B., & Egeth, H. E. (2002). Made you blink! Contingent attentional capture produces a spatial blink. *Attention, Perception & Psychophysics*, 64, 741-753.
- Folk, C. L., Leber, A. B., & Egeth, H. E. (2008). Top-down control settings and the attentional blink: Evidence for nonspatial contingent capture. *Visual Cognition*, 16, 616-642.
- Forster, S., & Lavie, N. (2008). Attentional capture by entirely irrelevant distractors. *Visual Cognition*, 16, 200-214.

- Forster, S. (2013). Distraction and mind-wandering under load. *Frontiers in Psychology*, 4(283), 1-6.
- Fox, E., Russo, R., & Dutton, K. (2002). Attentional bias for threat: Evidence for delayed disengagement from emotional faces. *Cognition & Emotion*, 16, 355-379.
- Glickman, M., & Lamy, D. (2018). Attentional capture by irrelevant emotional distractor faces is contingent on implicit attentional settings. *Cognition and Emotion*, 32(2), 303-314.
- Grimshaw, G. M., Kranz, L. S., Carmel, D., Moody, R. E., & Devue, C. (2018). Contrasting reactive and proactive control of emotional distraction. *Emotion*, 18(1), 26-38.
- Hahn, S., & Gronlund, S. D. (2007). Top-down guidance in visual search for facial expressions. *Psychonomic Bulletin & Review*, 14, 159-165.
- Hodsoll, S., Viding, E., & Lavie, N. (2011). Attentional capture by irrelevant emotional distractor faces. *Emotion*, 11(2), 346-353.
- Horstmann, G., Lipp, O. V., & Becker, S. I. (2012). Of toothy grins and angry snarls—Open mouth displays contribute to efficiency gains in search for emotional faces. *Journal of Vision*, 12(7), 1-15.
- IBM Corp. (2016). *IBM SPSS Statistics for Windows, Version 24.0*. Armonk, NY: IBM Corp.
- Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2, 194-203.
- Izard, C. E. (1994). Innate and universal facial expressions: Evidence from developmental and cross-cultural research. *Psychological Bulletin*, 115(2), 288-299.
- Jeffreys, H. (1961). *Theory of probability (3rd ed.)*. Oxford: Oxford University Press, Clarendon Press.
- Kennedy, B. L., Pearson, D., Sutton, D. J., Beesley, T., & Most, S. B. (2018). Spatiotemporal competition and task-relevance shape the spatial distribution of emotional interference

- during rapid visual processing: Evidence from gaze-contingent eye-tracking. *Attention, Perception, & Psychophysics*, 80(2), 426-438.
- Kennedy, B. L., Rawding, J., Most, S. B., & Hoffman, J. E. (2014). Emotion-induced blindness reflects competition at early and late processing stages: An ERP study. *Cognitive, Affective, & Behavioral Neuroscience*, 14(4), 1485-1498.
- Kohler, C. G., Turner, T., Stolar, N. M., Bilker, W. B., Brensinger, C. M., Gur, R. E., & Gur, R. C. (2004). Differences in facial expressions of four universal emotions. *Psychiatry Research*, 128(3), 235-244.
- Lamy, D. F., & Kristjánsson, Á. (2013). Is goal-directed attentional guidance just inter-trial priming? A review. *Journal of Vision*, 13, 1-19.
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). *International affective picture system (IAPS): Technical manual and affective ratings*. NIMH Center for the Study of Emotion and Attention, 39-58.
- Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, 9, 75-82.
- Le Doux, J. E. (1995). Emotion: Clues from the brain. *Annual Review of Psychology*, 46, 209.
- Le Doux, J. (1998). *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. New York: Simon and Schuster.
- Leblanc, É., Prime, D. J., & Jolicoeur, P. (2008). Tracking the location of visuospatial attention in a contingent capture paradigm. *Journal of Cognitive Neuroscience*, 20, 657-671.
- Lee, D. H., Susskind, J. M., & Anderson, A. K. (2013). Social transmission of the sensory benefits of eye widening in fear expressions. *Psychological Science*, 24(6), 957-965.
- Lichtenstein-Vidne, L., Henik, A., & Safadi, Z. (2012). Task relevance modulates processing of distracting emotional stimuli. *Cognition & Emotion*, 26, 42-52.

- Lipp, O. V., & Derakhshan, N. (2005). Attentional bias to pictures of fear-relevant animals in a dot-probe task. *Emotion*, 5, 365-369.
- LoBue, V., & Rakison, D. H. (2013). What we fear most: A developmental advantage for threat-relevant stimuli. *Developmental Review*, 33, 285-303.
- McHugo, M., Olatunji, B. O., & Zald, D. H. (2013). The emotional attentional blink: What we know so far. *Frontiers in Human Neuroscience*, 7, 151.
- Millisecond Software . (2016). *Millisecond Inquisit 5 [Computer Software]*. Seattle, WA :Millisecond Software .
- Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers*. 36, 630-633.
- Mogg, K., & Bradley, B. P. (2016). Anxiety and attention to threat: Cognitive mechanisms and treatment with attention bias modification. *Behaviour Research and Therapy*, 87, 76-108.
- Mohanty, A., & Sussman, T. J. (2013). Top-down modulation of attention by emotion. *Frontiers in Human Neuroscience*, 7(102), 1-5.
- Most, S. B., Chun, M. M., Widders, D. M., & Zald, D. H. (2005). Attentional rubbernecking: Cognitive control and personality in emotion-induced blindness. *Psychonomic Bulletin & Review*, 12, 654-661.
- Nako, R., Wu, R., Smith, T. J., & Eimer, M. (2014). Item and category-based attentional control during search for real-world objects: Can you find the pants among the pans? *Journal of Experimental Psychology: Human Perception and Performance*, 40, 1283-1288.
- Notebaert, L., Crombez, G., Van Damme, S., De Houwer, J., & Theeuwes, J. (2011). Signals of threat do not capture, but prioritize, attention: a conditioning approach. *Emotion*, 11, 81-89.

- Öhman, A. (1992). *Attention and information processing in infants and adults: Perspectives from human and animal research*. Hillsdale, NJ: Erlbaum.
- Öhman, A. (2005). The role of the amygdala in human fear: automatic detection of threat. *Psychoneuroendocrinology*, 30, 953-958.
- Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychological Review*, 108, 483-522.
- Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: detecting the snake in the grass. *Journal of Experimental Psychology: General*, 130, 466-478.
- Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, 42, 107-123.
- Peltola, M. J., Hietanen, J. K., Forssman, L., & Leppänen, J. M. (2013). The emergence and stability of the attentional bias to fearful faces in infancy. *Infancy*, 18, 905-926.
- Pergamin-Hight, L., Naim, R., Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., & Bar-Haim, Y. (2015). Content specificity of attention bias to threat in anxiety disorders: a meta-analysis. *Clinical Psychology Review*, 35, 10-18.
- Purkis, H. M., Lester, K. J., & Field, A. P. (2011). But what about the Empress of Racnoss? The allocation of attention to spiders and Doctor Who in a visual search task is predicted by fear and expertise. *Emotion*, 11, 1484-1488.
- Psychology Software Tools, Inc.. (2012). *E-Prime 2.0*. Retrieved from: <http://www.pstnet.com>.
- Raymond, J. E., & O'Brien, J. L. (2009). Selective visual attention and motivation: The consequences of value learning in an attentional blink task. *Psychological Science*, 20(8), 981-988.



- Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink?. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 849-860.
- Reeder, R. R., van Zoest, W., & Peelen, M. V. (2015). Involuntary attentional capture by task-irrelevant objects that match the search template for category detection in natural scenes. *Attention, Perception, & Psychophysics*, 77, 1070-1080.
- Reinecke, A., Rinck, M., & Becker, E. S. (2008). How preferential is the preferential encoding of threatening stimuli?: Working memory biases in specific anxiety and the Attentional Blink. *Journal of Anxiety Disorders*, 22, 655-670.
- RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA
- Schmidt, L. J., Belopolsky, A. V., & Theeuwes, J. (2015). Attentional capture by signals of threat. *Cognition and Emotion*, 29, 687-694.
- Singh, D., & Sunny, M. M. (2017). Emotion induced blindness is more sensitive to changes in arousal as compared to valence of the emotional distractor. *Frontiers in Psychology*, 8 (1381), 1-8.
- Smith, S. D., Most, S. B., Newsome, L. A., & Zald, D. H. (2006). An emotion-induced attentional blink elicited by aversively conditioned stimuli. *Emotion*, 6, 523-527.
- Soto, D., Hodsoll, J., Rotshtein, P., & Humphreys, G. W. (2008). Automatic guidance of attention from working memory. *Trends in Cognitive Sciences*, 12(9), 342-348.
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, 31(1), 137-149.
- Stein, T., Zwickel, J., Ritter, J., Kitzmantel, M., & Schneider, W. X. (2009). The effect of fearful faces on the attentional blink is task dependent. *Psychonomic Bulletin & Review*, 16, 104-109.

- Sui, J., & Humphreys, G. W. (2015). The integrative self: how self-reference integrates perception and memory. *Trends in Cognitive Sciences*, 19 (12), 719-728.
- Sussman, T. J., Jin, J., & Mohanty, A. (2016). Top-down and bottom-up factors in threat-related perception and attention in anxiety. *Biological Psychology*, 121, 160-172.
- Theeuwes, J. (1992). Perceptual selectivity for color and form. *Attention, Perception & Psychophysics*, 51, 599-606.
- Theeuwes, J. (1994). Endogenous and exogenous control of visual selection. *Perception*, 23, 429-440.
- Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*, 135, 77-99.
- Theeuwes, J. (2013). Feature-based attention: it is all bottom-up priming. *Philosophical Transactions of the Royal Society B*, 368, 1-11.
- Theeuwes, J., Reimann, B., & Mortier, K. (2006). Visual search for featural singletons: No top-down modulation, only bottom-up priming. *Visual Cognition*, 14(4-8), 466-489.
- Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, 32, 1349-1357.
- Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., ... & Nelson, C. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Research*, 168, 242-249.
- Turatto, M., & Galfano, G. (2000). Color, form and luminance capture attention in visual search. *Vision Research*, 40, 1639-1643.
- Van Dillen, L. F., Lakens, D., & Van Den Bos, K. (2011). At face value: Categorization goals modulate vigilance for angry faces. *Journal of Experimental Social Psychology*, 47(1), 235-240.

- Viechtbauer, W. (2010). Conducting meta-analyses in R with the Metafor package. *Journal of Statistical Software*, 36(3), 1-48.
- Vogt, J., De Houwer, J., Crombez, G., & Van Damme, S. (2013). Competing for attentional priority: Temporary goals versus threats. *Emotion*, 13(3), 587-598.
- Vromen, J. M., Lipp, O. V., Remington, R. W., & Becker, S. I. (2016). Threat captures attention, but not automatically: Top-down goals modulate attentional orienting to threat distractors. *Attention, Perception, & Psychophysics*, 78, 2266-2279.
- Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, 9, 585-594.
- Vuilleumier, P. (2015). Affective and motivational control of vision. *Current Opinion in Neurology*, 28(1), 29-35.
- Wells, A., & Matthews, G. (1994). *Attention and emotion: A clinical perspective*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Wyble, B., Folk, C., & Potter, M. C. (2013). Contingent attentional capture by conceptually relevant images. *Journal of Experimental Psychology: Human Perception and Performance*, 39, 861-871.
- Yantis, S. (1993). Stimulus-driven attentional capture. *Current Directions in Psychological Science*, 2, 156-161.
- Yiend, J. (2010). The effects of emotion on attention: A review of attentional processing of emotional information. *Cognition and Emotion*, 24, 3-47.
- Yokoyama, T., Padmala, S., & Pessoa, L. (2015). Reward learning and negative emotion during rapid attentional competition. *Frontiers in Psychology*, 6(269), 1-10.
- Zhang, J., & Mueller, S. T. (2005). A note on ROC analysis and non-parametric estimate of sensitivity. *Psychometrika*, 70(1), 203-212.

Zheng, C., Wang, J. Y., & Luo, F. (2015). Painful faces-induced attentional blink modulated by top-down and bottom-up mechanisms. *Frontiers in Psychology*, 6(695), 1-12.

Table 1. Means and standard errors for accuracy across all search goal and distractor conditions within Experiments 1 and 2. Data presented are arcsine transformed to account for significant skewness. The statistics for all pairwise comparisons between neutral distractors and their matched affective distractors are also reported, across both search goals. These statistics include *p*-values from pairwise *t*-tests, Hedges' *g* standardised effect sizes, Bayes factors and 95% confidence intervals. The Bayes factors were calculated using a half normal distribution with a mean of zero and an expected effect of 15% in Experiment 1, and 13% in Experiment 2.

Experiment	Search goal	Distractor condition	Mean (% accuracy)	<i>SE</i>	<i>p</i> -value	Hedges' <i>g</i>	<i>B<sub>H</sub></i>	95% CI [LB, UB]
Expt 1 ( <i>n</i> = 19)	Cute animals	Cute animal	52.93	4.11	< .001	.76	14896.33	[8.16, 17.81]
		Neutral	65.68	3.41				
		Threat	64.98	3.84	.738	.04	.19	[-4.02, 4.95]
	Threat animals	Cute animal	60.67	3.77	.876	.02	.19	[-2.86, 4.97]
		Neutral	61.05	3.83				
		Threat	51.18	3.58	.001	.60	224.34	[5.65, 14.82]
Expt 2 ( <i>n</i> = 18)	Cute animals	Happy face	56	3.94	.218	.16	.64	[-1.30, 7.23]
		Neutral face	58.86	4.49				
		Fearful face	56.80	4.39	.297	.11	.43	[-1.87, 5.86]
	Threat animals	Happy face	51.27	3.31	.232	.18	.61	[-1.41, 7.20]
		Neutral face	54.07	3.69				
		Fearful face	52.70	3.38	.389	.09	.28	[-1.54, 4.36]



Table 2. Mean arousal and valence ratings of target and distractor images by participants within Experiments 3a, 3b, 4, and 5 for stimuli included in their respective experiments. Standard deviations are presented in brackets. Ratings of each category represent the average of both distractor and targets together. The maximum positive valence was 9, whilst the maximum negative valence was 1, and 5 reflects neutral valence. The highest arousal rating was 9, whilst an arousal rating of 1 was reflects low arousal. All affective stimuli were significantly more arousing, and either more positive or negative than their neutral counterparts in the expected directions, all  $p$ 's < .005.

Stimulus category	Mean arousal ( <i>SD</i> )	Mean valence ( <i>SD</i> )
Threatening animals	6.50 (1.39)	3.08 (1.34)
Cute animals	3.97 (2.25)	7.93 (1.08)
Neutral animals	2.99 (1.28)	5.52 (1.19)
Fearful faces	4.72 (1.94)	3.19 (.98)
Happy faces	4.02 (1.74)	7.19 (1.33)
Neutral faces	2.60 (1.25)	5.10 (1.28)

Table 3. Means and standard errors for accuracy across all search goal and distractor conditions within experiments 3a and 3b. Data for Experiment 3a is untransformed, whilst data from Experiment 3b has been arcsine transformed to account for significant skewness. The statistics for all pairwise comparisons between neutral distractors and their matched affective distractors are also reported, across both search goals. These statistics include *p*-values from pairwise t-tests, Hedges' *g* standardised effect sizes, Bayes factors, and 95% confidence intervals. The Bayes factors were calculated using a half normal distribution with a mean of zero and an expected effect of 13%.

Experiment	Search goal	Distractor condition	Mean % accuracy	<i>SE</i>	<i>p</i> -values	Hedges' <i>g</i>	$B_{H[0, 13]}$	95% CI [LB, UB]
Expt 3a ( <i>n</i> = 18)	Cute animals	Fear face	51.62	4.42	.737	-.05	.17	[-5.64, 4.63]
		Neutral face	50.69	4.51				
		Threatening animal	50.69	4.61	.221	-.12	.07	[-5.92, 1.31]
		Neutral animal	48.38	4.21				
	Threat animals	Fear face	53.94	3.68	.787	-.04	.17	[-5.70, 3.70]
		Neutral face	53.24	3.99				
		Threatening animal	47.92	3.30	< .001	.69	$109 \times 10^3$	[7.18, 15.68]
		Neutral animal	59.49	4.40				
Expt 3b ( <i>n</i> = 18)	Cute animals	Fear face	69.46	4.35	.549	< .01	.12	[-5.66, 2.82]
		Neutral face	67.99	3.88				
		Threatening animal	64.93	3.87	.998	-.08	.23	[-5.17, 5.36]
		Neutral animal	64.85	3.62				
	Threat animals	Fear face	58.83	3.57	.986	< .01	.24	[-5.46, 5.68]
		Neutral face	59.09	4.16				
		Threatening animal	32.92	2.39	< .001	2.46	$925 \times 10^8$	[23.58, 38.86]
		Neutral animal	64.16	3.36				



Table 4. Means and standard errors for  $A'$  across all search goal and distractor conditions within Experiments 4 and 5. Data for Experiment 4 is untransformed, whilst data from Experiment 5 has been arcsine transformed to account for significant skewness. The statistics for all pairwise comparisons between neutral distractors and their matched affective distractors are also reported, across both search goals. These statistics include  $p$ -values from pairwise t-tests, Hedges'  $g$  standardised effect sizes, Bayes factors, and 95% confidence intervals. The Bayes factors were calculated using a half normal distribution with a mean of zero and an expected effect of .11  $A'$ .

Experiment	Search goal	Distractor condition	Mean $A'$	$SE$	$p$ -values	Hedges' $g$	$B_{H[0, .11]}$	95% CI [LB, UB]
Expt 4 ( $n = 18$ )	Happy faces	Fear face	.88	.02	.044	.71	3.67	[1.39, 12.72]
		Neutral face	.93	.01				
		Threatening animal	.92	.01	.467	.14	.21	[-.01, .03]
		Neutral animal	.93	.01				
	Fearful faces	Fear face	.84	.03	.001	.80	123.70	[5.61, 16.11]
		Neutral face	.92	.02				
		Threatening animal	.93	.02	.171	.38	.68	[< .01, .06]
		Neutral animal	.96	.01				
Expt 5 ( $n = 25$ )	Cute animals	Fear face	.84	.03	.777	.01	.19	[-.03, .04]
		Neutral face	.84	.04				
		Threatening animal	.85	.03	.093	-.14	.05	[-.05, <.01]
		Neutral animal	.82	.03				
	Threat animals	Fear face	.91	.02	.837	.09	.31	[-.04, .03]
		Neutral face	.92	.02				
		Threatening animal	.87	.03	.028	.25	3.42	[.01, .07]
		Neutral animal	.90	.02				

### Figure captions

Figure 1. Structure of a single RSVP trial and example stimuli from across six Experiments; nine images were presented per trial for either 83 (in Experiment 4) or 100ms (in Experiments 1, 2, 3a, 3b, and 5) per image with no inter-stimulus interval. Examples of the instruction frame depict how the search instructions were presented in Experiments 1, 2, 3a, 3b, and 5, Experiment 4 differed by instructing participants to search for happy and scared faces. The distractor and target frame examples depict the categories of images which appeared as stimuli, as well as the location of the distractor in each task, as symbolised by dashed lines. In Experiments 1, 2, 3a, 3b, and 5, threatening and cute animal targets were presented as targets, whilst in Experiment 4 happy and fearful face targets were presented. At the end of the trial participants were either required to type out what the specific animal target had been (Experiments 1-3b), or whether the target had been present or absent from the RSVP stream (Experiment 4 and 5). The emotional faces in the figure were taken from the NimStim image set with permission for publication (Tottenham et al. 2009).

Figure 2a and 2b. Affective distractor effects (% correct neutral distractor - % correct affective distractor) for cute and threatening animal distractors across both cute and threatening animal search conditions in Experiments 1 and 2. The error bars represent within-subjects standard error, and asterisks denote a significant difference from the match neutral distractor:  $p < .05 = *$ ;  $p < .001 = **$ .

Figure 3a and 3b. Affective distractor effects (% correct neutral distractor - % correct threat distractor) for fearful face and threatening animal distractors across both cute and threatening animal search conditions in Experiment 3a and Experiment 3b. The error bars represent

within-subjects standard error, and asterisks denote a significant difference from the match neutral distractor:  $p < .05 = *$ ;  $p < .001 = **$ .

Figure 4. Affective distractor effects ( $A'$  neutral distractor –  $A'$  threat distractor) for fearful face and threatening animal distractors across both happy and fearful face search conditions in Experiment 4. The error bars represent within-subjects standard error, and asterisks denote a significant difference from the match neutral distractor:  $p < .05 = *$ ;  $p < .001 = **$ .

Figure 5. Affective distractor effects ( $A'$  neutral distractor –  $A'$  threat distractor) for fearful face and threatening animal distractors across both happy and fearful face search conditions in Experiment 5. The error bars represent within-subjects standard error, and asterisks denote a significant difference from the match neutral distractor:  $p < .05 = *$ ;  $p < .001 = **$ .

Figure 1.

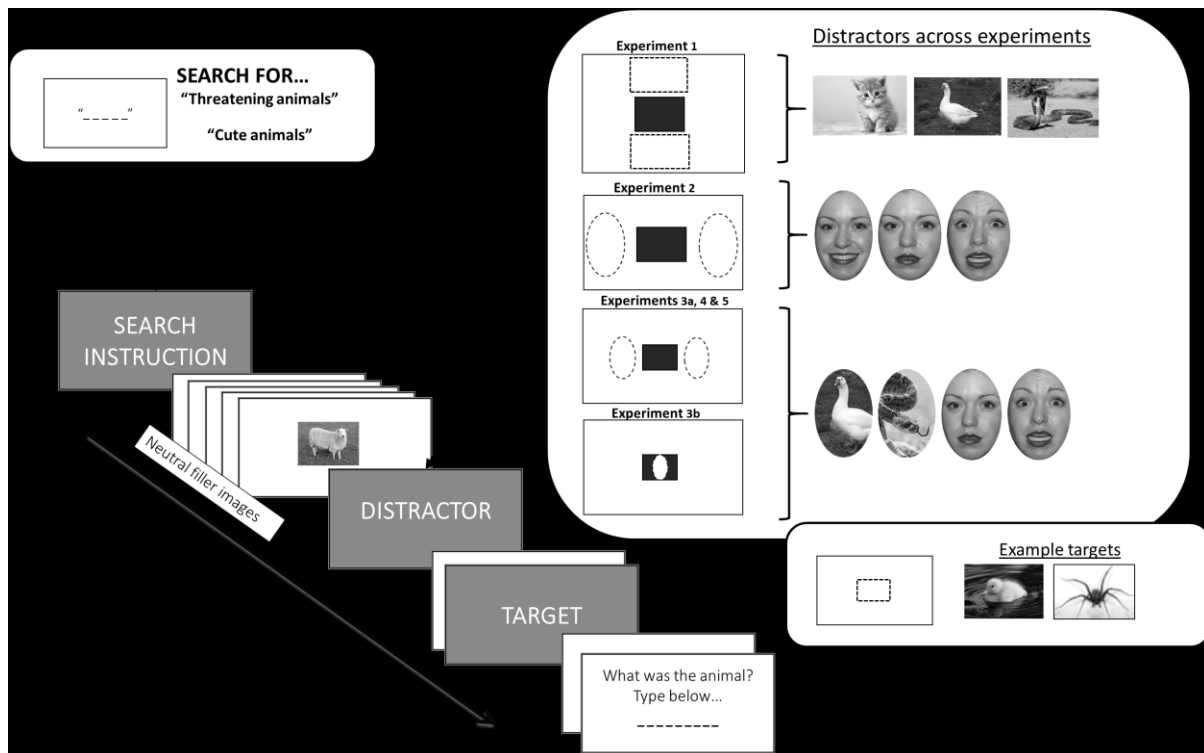


Figure 2a and 2b.

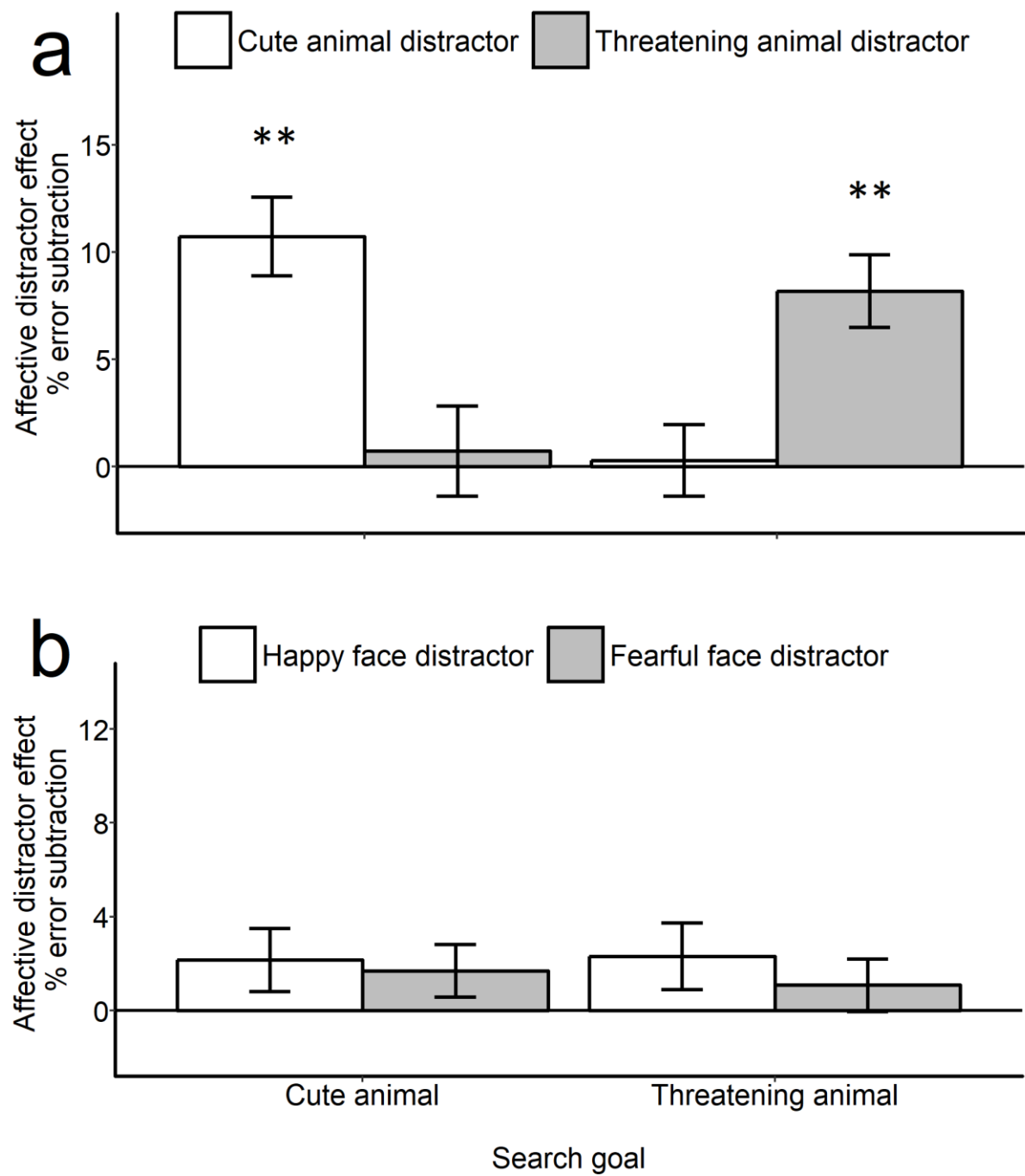


Figure 3a. and 3b

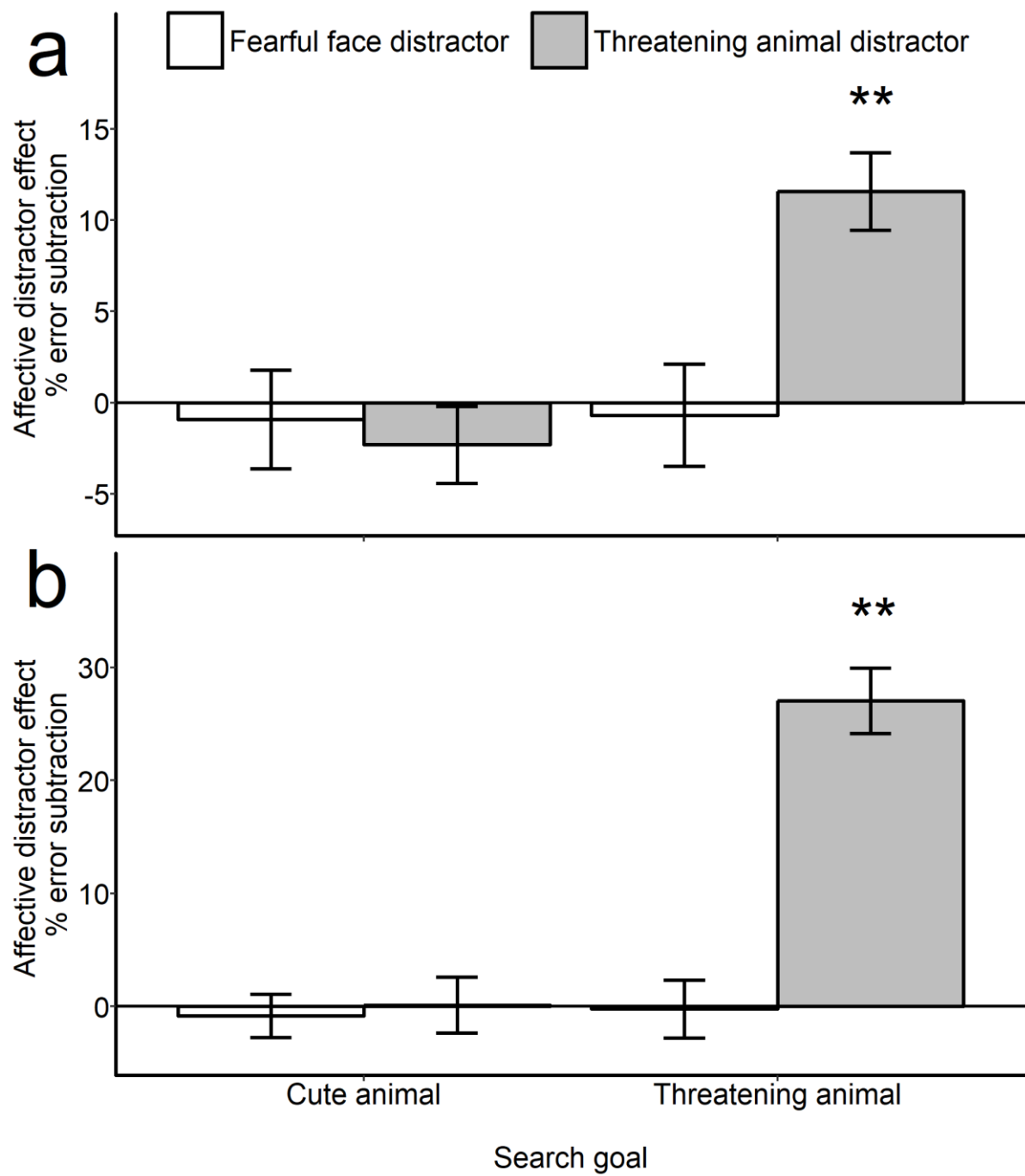


Figure 4.

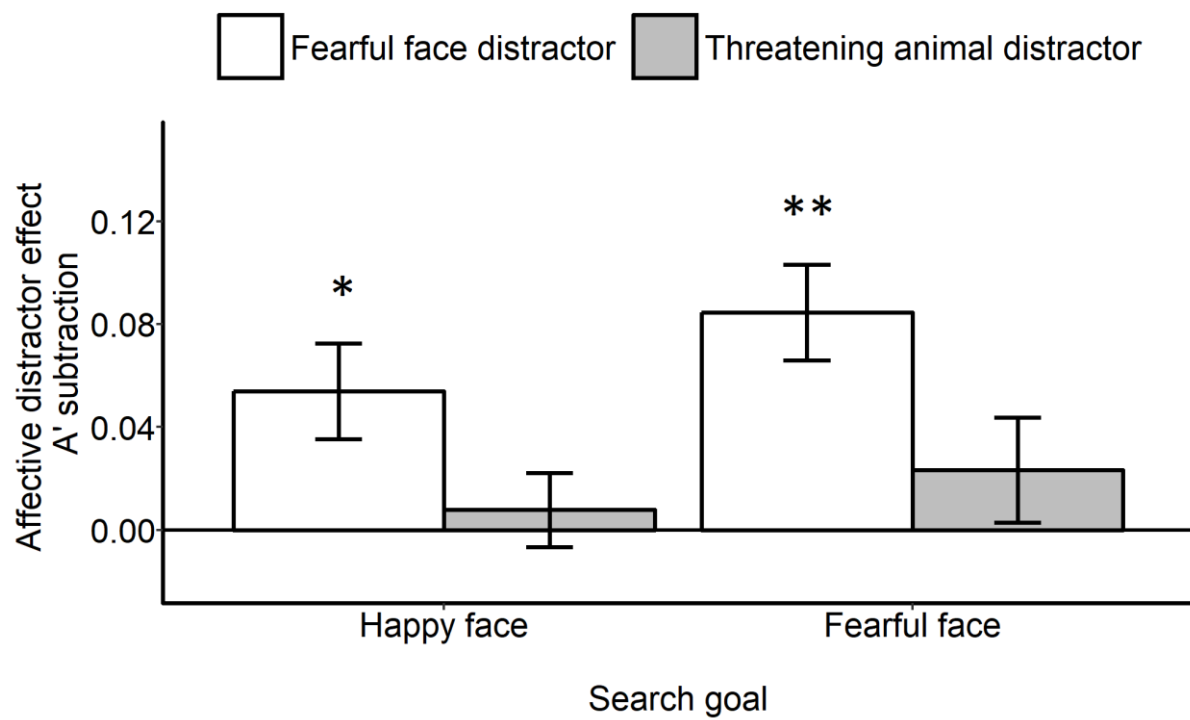


Figure 5.

